

### Von der Rechenmaschine zum "neurotischen Computer"?

Hinkelmann, Reinhold; Paluek, J. Ute

Verffentlichungsversion / Published Version  
Zeitschriftenartikel / journal article

#### Empfohlene Zitierung / Suggested Citation:

Hinkelmann, R., & Paluek, J. U. (1991). Von der Rechenmaschine zum "neurotischen Computer"? *Psychologie und Gesellschaftskritik*, 15(2), 21-38. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-290439>

#### Nutzungsbedingungen:

Dieser Text wird unter einer CC BY-NC-ND Lizenz (Namensnennung-Nicht-kommerziell-Keine Bearbeitung) zur Verfgung gestellt. Nhere Ausknfte zu den CC-Lizenzen finden Sie hier:  
<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.de>

#### Terms of use:

This document is made available under a CC BY-NC-ND Licence (Attribution-Non Comercial-NoDerivatives). For more Information see:  
<https://creativecommons.org/licenses/by-nc-nd/4.0>

Reinhold Hinkelmann & J. Ute Palußek

## VON DER RECHENMASCHINE ZUM "NEUROTISCHEN COMPUTER"?

### 1.

Grundsätzlich verstehen wir Computer als vom Menschen geschaffene, veräußerlichte, als Maschinen objektivierte Denkmodelle, die den "klassischen Maschinenbegriff" entgrenzen. Computer haben gegenüber bisherigen Maschinen eine neue Qualität erreicht. Während bei der "klassischen Maschine" Inhalt und Gestalt zusammenfallen, sie also ihr Inhalt *ist* (diesen verkörpert), existiert die neue, "transklassische Maschine" zunächst inhaltsleer. Die verschiedensten Gebiete können unter bestimmten Voraussetzungen bearbeitet werden, sie *hat* also einen Inhalt (bekommt diesen nachträglich). Diese Kontextungebundenheit oder "Inhaltsleere" kennzeichnet die "transklassische Maschine" als Verkörperung der abendländischen abstrakten Logik. Der Computer ist "implementierte Theorie" (Holling & Kempin 1989). Die Implementierungsebene bezeichnet "die technisch-konkrete Umsetzung eines Algorithmus<sup>1</sup> in einer bestimmten Computerarchitektur und Programmiersprache" (Strube 1990, S. 137).

Weil der Computer implementierte Theorie ist, seine Programme aus Algorithmen bestehen, die am Verhalten der Maschine (output) orientiert sind, realisiert "er" programmspezifische Handlungstheorien. Z. B. steuern die Rechengvorgänge eines computergestützten Verkehrsleitsystems unmittelbar die Ampelschaltungen. Zwischen Denken und Handeln gibt es beim Computer keinen Hiatus mehr. Die Frage "Denkt (und handelt) der Computer?" ist keine mehr. Mensch und "transklassische Maschine" teilen die über Sozialisation bzw. Programmierung implementierte abendländische Logik

<sup>1</sup> "Algorithmen sind (in eine formale Struktur gebrachte; d. Verf.) detaillierte Verhaltensmuster zur automatischen Lösung von Problemen" (Bammé u. a. 1983, S. 115).

(und deren Handlungsrelevanz). Dem entspricht die These von Bammé u. a. (1983), daß die Schnittstelle zwischen Mensch und Maschine durch das Subjekt verläuft und nicht, wie üblicherweise behauptet, an der (Körper-) Peripherie von Mensch und Maschine (lokalisiert ist).

## 2. Zweimal künstliche Intelligenz - und ihre Ursprünge

Das Projekt der Künstlichen-Intelligenz-Forschung ist zunächst einmal der Versuch, unterschiedliche mathematische Formalismen in Algorithmen umzusetzen: Was mit Hilfe von Formalismen "überhaupt berechenbar ist, so die unbewiesene, aber auch seit Jahrzehnten unwiderlegte These, läßt sich durch jedes dieser formalen Mittel berechnen" (Strube 1990, S. 137). Potentiell kann jeder mathematische Formalismus (ein Beispiel ist die "Turing-Maschine") in Algorithmen umgesetzt werden, wobei die technische Realisierung (Implementierung) bisher nur für den digitalen Computer erfolgte.

Wir wollen im folgenden zwei theoretische Modelle (Formalismen) betrachten, die in der KI-Forschung aufgenommen (und manchmal auch gegeneinander ausgespielt) wurden: die *informationsverarbeitende KI* und die, keineswegs neue, aber wiederentdeckte *wachsende KI*. Beide Bezeichnungen gebraucht Turkle (1984; 1988) für zwei Formen der unterschiedlichen theoretischen Annäherung an das Problem der Simulation menschlicher Handlungen durch Computer. Die Begriffe, mit denen Turkle die beiden Modelle unterscheidet, verschleiern mehr, als sie klären: auch die *wachsende KI* verarbeitet Informationen und bekommt über die Bezeichnung "wachsend" eine mythologische Dimension (vgl. Turkles euphorische Beschreibung; 1984, S. 348 ff.; 1988). Wir verwenden deshalb im weiteren Text statt informationsverarbeitende KI den Begriff *Symbolverarbeitung* und statt wachsender KI den Begriff *Neokonnektionismus* (vgl. Strube 1990).

### *Was sind die Unterschiede?*

Computermodelle nach Art der Symbolverarbeitung sind die digitalen Computer, die wir kennen, mit denen wir bereits arbeiten. Entsprechend nennt Turkle (1988) diese Richtung auch "normalisierende" KI. Ihre theoretischen Grundlagen sind an der zweiwertigen (Prädikaten-)Logik (alternative Zustände: ja/nein) orientiert, ihre Programme enthalten Regeln, ihre Anwendung erfolgt deduktiv auf ein spezifisches Problem. Der Computer arbeitet monokausal (wenn a, dann b), und die Programmabschnitte laufen seriell (nacheinander) ab. Ihre Anwendung finden sie z. B. in Expertensystemen oder Decision Support Systems. Ihr Einfluß auf die Kognitive Psychologie wird z. B. bei Norman (s. u.) sichtbar.

Die Modelle des Neokonnektionismus sind bis jetzt vor allem Theorie geblieben, ohne eine technische Umsetzung in entsprechend konstruierte Computer. Diese Richtung trägt für Turkle "subversiven" Charakter, doch darauf kommen wir später zu sprechen. Die theoretischen Grundlagen sind hier biologisch, orientiert an neuronalen Strukturen, wenngleich diese wenig Ähnlichkeit mit aktuellen Neuronenmodellen der Neurophysiologie haben. Ihr Grundelement ist demnach ein Formal-Neuron, das mit vielen anderen Formal-Neuronen verknüpft ist (Vernetzung). Ihre Grundlagen sind an der Stochastik, also an der Gewichtung von Parametern orientiert. Der (hypothetische) Computer dieses Modells folgert induktiv, er lernt aus Beispielen, aus Erfahrung. In diesem neokonnektionistischen Modell werden Informationen durch den Zusammenhang (Zustand der Vernetzung) von Formal-Neuronen (oder "Agenten" in Analogie zu einem gesellschaftstheoretischen Demokratiemodell bei Minsky) verarbeitet. Deshalb spricht man auch von einer "subsymbolischen" Verarbeitung von Informationen. Dabei werden die Informationen simultan bearbeitet, ihre Reihenfolge der Eingabe ist bedeutsam, sie sind kontextabhängig.

Als Vergleich greifen wir das Beispiel der Wettervorhersage nach Turkle auf. Ein symbolverarbeitender Computer arbeitet mit einem Programm, das aus Regeln besteht, aus denen er ableitet, ob bei gegebenen Informationen Regen vorausgesagt werden kann oder nicht. So ein Computerprogramm zur Prognose funktioniert wie *ein* Meteorologe. Im neokonnektionistischen

Modell gehen dieselben Informationen an eine ganze Reihe von "Agenten", die alle nach ihren je eigenen Regeln eine Vorhersage treffen. Der Computer in diesem Modell macht dann sozusagen ein Meinungsbild wie bei einer *Gruppe* von Meteorologen, wobei die vergangene Treffsicherheit der einzelnen als Gewichtung berücksichtigt wird. Wenn ein Agent hier ausfällt oder falsch urteilt, kann das gesamte Programm dennoch brauchbare Aussagen liefern.

### 3. KI-Forschung und Psychologie

Die grundlegende These, von der aus die KI-Forschung ihre theoretischen Konzepte entwickelt, könnte nach Turkle (1988, S. 241) lauten: "... die Essenz (das Wesen) geistigen Lebens kann als "set" von Regeln (Prinzipien) beschrieben werden, das auf Menschen und Maschinen gleichermaßen zutreffen kann". Die methodologische Prämisse dabei "is that if one builds a machine that can do something intelligent, the way one gets the machine to do it is relevant to thinking about how people do it as well" (Turkle 1988, S. 244).

Die Aufgabe der Programmentwicklung innerhalb des Systems der formalen Logik (Satz der Identität - Satz vom Widerspruch - Satz vom ausgeschlossenen Dritten), mit Hilfe von Algorithmen bietet natürlich auch eine Chance über menschliches Denken (Intelligenz) und Handeln nachzusehen. In den 60er Jahren half die Computerwissenschaft der Psychologie die "Black Box" des Behaviorismus wiederaufzufüllen. So habe die "Existenz" des Computers die Gedächtnisforschung und die der "inneren Zustände" innerhalb der akademischen Psychologie "(re)legitimiert" und den dazugehörigen Theorien (Kybernetik und Automaten-Theorie) zu größerer Gewichtung verholfen (KI als "stützender Mythos" (sustraining myth); Turkle 1988).

Die KI-Forschung nach dem Prinzip der Symbolverarbeitung gab in den 60er Jahren der Allgemeinen Psychologie wesentliche Impulse. "Die Einführung dieses Kriteriums (der KI, ob und wie gut ein Programm die verlangten Leistungen erbringt; d. Verf.) hat ungeheuer befreiend gewirkt;

man durfte, ja man mußte von nun an Modelle ersinnen, die der Komplexität des Gegenstandes gerecht werden. Dies hat zu einer Lebendigkeit des Theoretisierens geführt. ... Es ist eine methodische und thematische Liberalisierung eingetreten ..." (Neumann 1985, S. 11). Die Kognitive Psychologie (Neisser 1967; dt. 1974) formierte sich, ihre Zielrichtung war und ist eine Handlungstheorie in einem "Modell der hierarchisch-sequentiellen Handlungsorganisation", so der Titel eines Beitrags von Volpert (1982).

Interessant an dieser Entwicklung einer Wissenschaft war, daß die Psychologen dieses Fachgebietes bald das Unbefriedigende des ausschließlich regelgeleiteten Handlungsmodells bemerkten und zusätzliche Faktoren in ihr Modell aufnahmen, wobei sie die Grundannahmen beibehielten. So beklagt z. B. Dörner die einseitige Betonung "epistemischer" und die Vernachlässigung "heuristischer" Kompetenzen. "Die epistemische Kompetenz beinhaltet die Abschätzung des Individuums, mit der Situation aufgrund seiner *Erfahrungen*, etwa aufgrund seines Handlungswissens fertig werden zu können. Die *heuristische Kompetenz* beinhaltet die Abschätzung der Fähigkeit, neuartige Situationen, für die Wissen nicht vorliegt, bewältigen zu können" (Dörner 1982, S. 32). Er beklagt das Fehlen der Zielaufarbeitung, der Selbstreflexion, der Sammlung und Strukturierung von Information, des Gesamtgefüges der Absichten und der emotionalen Einbettung des Denkablaufs in der Operator doktrin vom Ablauf problemlösenden Denkens; d. h. er vermißt neben der regelanwendenden die regelgenerierende Komponente problemlösenden Denkens.

Soweit die symbolverarbeitende KI. Das neokonnektionistische Modell versucht in etwas dem nachzukommen, was Dörner fordert. Hier geht es um Heuristiken, um Lernen aus Erfahrungen und Beispielen. Das "Agentenmodell" beschreibt, wie Intelligenz über die Interaktion einer Gesellschaft von Agenten erwachsen (emerge) soll. Die interagierenden Agenten sind in ihrer Funktion "beschränkt" (also "dumm"), ihre simultane Aktivierung in der Vernetzung produziert neues Wissen, das über die Summe des Wissens der einzelnen Agenten hinausgeht.

2. Der Mensch ist ein informationsverarbeitendes System. "Als theoretischer Psychologe möchte ich formale der menschlichen Informationsverarbeitenden Strukturen entwickeln" (ebd., S. 44).
3. "X zu verstehen bedeutet ein Computerprogramm zu erstellen, das X simuliert." Die von Norman entwickelte Handlungstheorie ist wie ein Computerprogramm mit Haupt- und Unterprogrammen aufgebaut.

Wie also versteht Norman eine einfache Handlung, wie z. B. das Bedienen einer Kaffeemaschine?

Am Anfang steht die Absicht, "Kaffee trinken zu wollen", die sich möglicherweise daraus ableiten läßt, daß der vorliegende Artikel bisher so trocken zu lesen war. Ist die Absicht einmal gefaßt, so schaltet sie ein spezifisches Hauptprogramm (parent schema) ein, das wiederum automatisch die zur Ausführung der Absicht notwendigen Unterprogramme (child schemes) aktiviert.

Konkret heißt das dann: Wasser in die Kaffeemaschine einfüllen, Filter und Kaffee an den richtigen Ort bringen und die Maschine einschalten. Jedem parent schema ist dabei eine begrenzte Zahl von child schemes zugeordnet, die zur Realisierung der Absicht nützlich sind. Will man/frau Kaffee machen, so genügt es in der Regel, die Absicht zu fassen, alles Weitere läuft dann automatisch ab. Das Herausbilden der Absicht löst außerdem bestimmte Kontrollprozesse aus, die festlegen, wann welche motorischen Aktivitäten wie in welcher Reihenfolge ablaufen sollen, um nicht das Wasser in der Kanne zu lassen, den Kaffee in den Wasserbehälter zu füllen und den Filter in den Papierkorb zu werfen.

Sollte etwas schiefgehen, so handelt es sich nach Norman um einen Fehler/Irrtum (error). Er verwendet die Begriffe Fehler/Irrtum entsprechend dem Alltagsverständnis im Sinne eines sinnlosen und zufälligen Geschehens. Die motorischen Aktivitäten bei der Ausführung einer Absicht laufen meist unter der Bewußtseinsschwelle ab, also ohne bewußte Steuerung. Diese Handlungen "sind unabhängige, strukturierte Zusammensetzungen von Kontrollprozessen oder 'schemes'. Sie werden auf verschiedene Weise in Bereitschaft geschaltet oder 'aktiviert': durch das Kurzzeitgedächtnis,

durch Motive oder durch Hinweisreize aus der Situation (Norman, 1980, S. 44). Für Norman sind die Mechanismen des Systems, die seine Stärken und Leistungsfähigkeit ausmachen, zugleich für das Auftauchen von Fehlern verantwortlich, und eben diese menschlichen Fehler bezeichnet er als Fehlleistungen. "Eine Fehlleistung ist eine Form menschlichen Irrtums, definiert als die Ausführung einer Handlung, die nicht intendiert ist" (1981, S. 1).

Grundsätzlich kategorisiert Norman die Fehlleistungen "je nachdem, welcher Teil der menschlichen Maschine beteiligt ist" (1980, S. 44) und führt sie auf drei Ursachen zurück:

1. Wenn das Entstehen einer Intention irrtümlich verläuft. Eigentlich will man/frau, um bei unserem Beispiel zu bleiben, lieber Tee als Kaffee trinken.
2. Durch falsches Aktivieren von Schemen. Wenn man/frau auf dem Nachhausewege von der Arbeit noch eine Kleinigkeit einkaufen will, aber wie gewohnt auf direktem Weg nach Hause fährt.
3. Durch Fehler im Auslösen von Schemen. Entweder wird die Handlung zum falschen Zeitpunkt oder überhaupt nicht ausgelöst. Man/frau will jemanden um einen Gefallen bitten und stellt nach geraumer Zeit fest, daß die Bitte nicht erfüllt wurde. Auf beleidigtes Nachfragen erfährt man/frau dann, daß man/frau die Bitte nie vorgetragen hat.

Zusammenfassend lassen sich bei Norman Fehlleistungen als Irrtümer in der symbolverarbeitenden Maschine Mensch begreifen. Irrtümer/Fehler sind dabei als sinnlose, dysfunktionale und zufällig auftretende Elemente menschlichen Handelns zu verstehen. Der Begriff Fehlleistung hat somit einen wichtigen Teil seines Inhalts verloren, den der *Leistung*.

... und Freud?

Wie steht Norman zum Ansatz von Freud, dessen Begriff der Fehlleistung er übernimmt? Im Titel seines Aufsatzes behauptet er ja ein post-feudianisches Modell für Fehlleistungen gefunden zu haben, also Freud überholt und dessen alte Theorie hinter sich gelassen zu haben.



Ein Beispiel einer Fehlleistung, das gleichermaßen von Norman und Freud behandelt wurde:

"Ein Präsident unseres Abgeordnetenhauses eröffnete einmal die Sitzung mit den Worten: Meine Herren, ich konstatiere die Anwesenheit von ... Mitgliedern und erkläre somit die Sitzung für *geschlossen*" (Freud 1915/16, S. 57).

Für Norman sind ausschließlich die formalen Gesichtspunkte für das Zustandekommen einer Fehlleistung relevant. Er mißdeutet Freud, wenn er ausführt, daß "diese Handlung (des Präsidenten; d. Verf.) verborgene Motive zu enthüllen" (1980, S. 43) scheint, aber es vor allem wichtig ist, daß "die versteckte Absicht eine ziemlich präzise Gelegenheit (benötigt), um sich zu entfalten" (ebd.). Diese wiederum hängt ab von syntaktischen, phonologischen u.a. Auslösefaktoren. Im betreffenden Beispiel kann nur das eine Wort im Satz ("eröffnet") durch das geäußerte andere ("geschlossen") ersetzt werden, ohne die Regeln der Grammatik zu verletzen. Norman interessiert sich nicht dafür, daß bei dem Versprecher des Präsidenten ein inhaltlich sinnvoller und verständlicher Satz gebildet wird. Er berücksichtigt lediglich *eine* Intention, deren "Fehlgehen" er auf die oben beschriebenen Mechanismen zurückführt. Für ihn ist die Art und Weise, also das *Wie* man/frau sich verspricht, die Antwort auf die Frage, *Warum* man/frau sich verspricht. Demgegenüber Freud: "Im vorliegenden Falle wird die Erklärung wohl diese sein, daß der Präsident sich wünschte, er wäre schon in der Lage, die Sitzung, von der wenig Gutes zu erwarten stand, zu schließen, aber - eine häufige Erscheinung - der Nebengedanke setzte sich wenigstens teilweise durch und das Resultat war >geschlossen< für >eröffnet<, also das Gegenteil dessen, was zu sprechen beabsichtigt war" (1901, S. 54). Der Präsident wurde demnach von seinen eigenen Worten überrascht, deren Inhalt er offensichtlich nicht kontrollieren konnte.

### *Widersprüche*

In diesem kurzen Vergleich beider Ansätze hat es den Anschein, als ob Norman Freud mißverstehen muß, um seine theoretische Position zu retten. Indem er Freud den menschlichen Bereich der "geheimen Wünsche und Absichten" zuschreibt, kann er sich auf scheinbar "einfachere Erklärungen"

konzentrieren. Übrig bleibt dabei lediglich der *Mechanismus* der Fehlleistungen: Norman (1980, S. 44) meint, "...daß Fehlleistungen durch Widerstreit und Vermischen von zugrundeliegenden seelischen Verarbeitungsmechanismen entstehen, die oft parallel zueinander ablaufen".

Den Einwand, daß es sich bei den Fehlleistungen "um kleine Entgleisungen der Funktion, Ungenauigkeiten des Seelischen, deren Bedingungen sich angeben ließen" handelt, kannte bereits Freud (1915/16, S. 52). Er bestritt auch nicht die Wichtigkeit dieser Mechanismen, vielmehr betonte er selbst: "Die Lauteinflüsse, Wortähnlichkeiten und die von den Worten auslaufenden gebräuchlichen Assoziationen sind wiederum als bedeutsam anzuerkennen. Sie erleichtern das Versprechen, indem sie ihm Wege weisen, die es wandeln kann. Aber wenn ich einen Weg vor mir habe, ist damit auch wie selbstverständlich entschieden, daß ich ihn gehen werde?" (1915/16, S. 68). Die Wahl des Weges, den der Wanderer gehen wir, ist kein passives Geschehen, das zufällig eintritt, sondern der Niederschlag der *Absichten* des Wanderers, also ein aktiver Entscheidungsprozeß. Bei dem Versprecher des Parlamentspräsidenten macht der zustandegekommene Satz zugleich einen *Sinn*. Aus den Begleitumständen der (psychischen) Situation des Sprechers läßt sich erschließen, wie bei ihm zwei Absichten - Pflicht und Neigung - in Widerspruch zueinander geraten. Das Versprechen stellt eine Lösung des Konflikts dar, eine Kompromißlösung, die ein halbes Gelingen und ein halbes Mißlingen jeder Absicht impliziert.

Es können also zwei Absichten unterschieden werden: einmal die bewußte und eine zweite, die diese stört. Die störende Absicht kann dabei in einer inhaltlichen Beziehung zur gestörten stehen (ihr widersprechen, sie berichtigen oder ergänzen) oder aber beide Absichten beziehen sich auf unterschiedliche Assoziationsbereiche, haben also keine inhaltliche Beziehung zueinander.

### *Mensch = Maschine*

Norman knüpft an seine Vorgehensweise einen technologischen Anspruch, wenn er sagt, "wenn wir wissen, wie sich Fehler ereignen, dann kann es

auch möglich sein, sie zu verhindern" (1980, S. 44). D. h. man/frau muß an den Stellen ansetzen, an denen die Maschine außer Kontrolle gerät und lediglich das fehlerhafte Maschinensystem vervollkommen. Das Entscheidende aber, das mit einer Fehlleistung einhergeht, ist der *Kontrollverlust* der Person, die nicht im Sinne ihrer bewußten Absicht handelt. Diese Kränkung, die durch das Erleben des Kontrollverlustes entsteht (daß wir manchmal Handlungen ausführen, die wir so nicht machen wollten), will uns Norman ersparen. Und zwar indem er uns erstens in der Funktionsweise mit der Maschine (Computer) gleichsetzt und uns zweitens das Bild der "vollkommenen Maschine" ausmalt, dem wir uns durch verstärkte technische Eingriffe zunehmend nähern können. Der Effekt besteht in der Aufhebung des Unterschiedes zwischen Mensch und Maschine. Was jenseits der Schnittstelle im Menschen als unreduzierbarer Rest bestehen bleibt, verschwindet. Übrig bleibt die Identität (Mensch=Maschine).

Für den Maschinen-Menschen ist der Kontrollverlust kein Problem mehr, da Irrtümer und slips der Tendenz nach durch noch mehr Maschine bzw. Technik eliminiert werden können. Wenn wir uns einen solchen Maschinen-Menschen vorstellen, so hat er das genuin Menschliche verloren.

Das spezifisch Menschliche an dieser Stelle ist der Widerspruch zwischen bewußter und störender Absicht. Diesen Widerspruch macht Norman zum Fehler und zerzt ihn aus seiner dunklen Unberechenbarkeit in den hellen Schein der Beherrschbarkeit. Dabei verliert der Mensch seine unbekannte, störende Absicht; übrig bleibt eine fehlerhafte, kaputte Einheit. Ein defekter Mechanismus vermag nicht die eine Absicht gegen eine andere, willentliche zu stellen und kann darum auch keine Konfliktlösung hervorbringen. Diese Konfliktlösung stellt - im psychoanalytischen Sinne - in erster Linie eine kreative Leistung dar, weil ja etwas anderes, höchst Sinnvolles dabei herauskommt. Bei Norman bleibt der Verlust von Kontrolle immer nur ein Fehlen, ein Zuwenig, das nicht an die "richtige" Handlung heranreicht. Jede Abweichung ist störend.

Norman zieht die Kränkung vom Kontrollverlust ab, indem er die Funktionsweise des Menschen mit dem Ideal der perfekten Maschine gleichsetzt. Er entzaubert so die "Störungen" von ihrem Sinn und ihrer Bedeutung im

Handeln des Einzelnen. Dies reduziert jegliche Störung auf Dysfunktionalität der Maschine Mensch. Diese wiederum läßt sich nur durch ein Mehr an Funktion beheben, ausgleichen oder kompensieren.

#### 4.2 Das neokonnektionistische Modell

Auch innerhalb der KI wird die symbolverarbeitende KI wegen ihrer mechanistischen Denkweise und ihrer monokausalen Erklärungsversuche (vgl. Turkle 1988, S. 249ff.) kritisiert. Jedoch bietet sich mit der neokonnektionistischen KI ein (Denk-)Modell an, das der psychoanalytischen Auffassung von der "Überdeterminiertheit" psychischer Phänomene eher entspreche. Schon immer hätte es "traditionelle Affinitäten" zwischen Neokonnektionismus und Psychoanalyse gegeben, so z. B. die Betonung der Selbstreflexion. Wie jener beschäftigte sich auch die psychoanalytische Objektbeziehungstheorie mit "inneren Objekten", unbewußten Suborganisationen des Ich (z. B. das Über-Ich), die über Introjektionen und Identifikationen "nach innen gebracht" werden. In der Orientierung an "inneren Objekten" manifestiert sich auch eine gemeinsame "subversive Kraft" der beiden Richtungen, da die Interaktion der Objekte den "output" präge und somit die Vorstellung vom "autonomen Ich" auflöse. Das Ich könne nur noch fragmentiert gedacht werden. Seine "Stimme" (output/Bewußtsein) ist das Resultat einer "cacophony of voices" der Agenten/inneren Agenten. Das "autonome Ich" ist zum "dezentralen Selbst" geworden.

Sherry Turkle parallelisiert die Entwicklung innerhalb der psychoanalytischen Theorie und der KI unter dem Stichwort des "Paradigmenwechsels" (Kuhn) vom Trieb- zum Objektmodell bzw. vom Symbol zum Agenten (Subsymbol):

"Fairbairn ersetzte die Freudschen Dichotomien von Ich und Es, Struktur und Energie mit unabhängigen Agenturen innerhalb der Psyche, die denken, wünschen und Bedeutungen erzeugen dadurch, daß sie interagieren, ebenso wie bei der wachsenden KI autonome Agenten innerhalb des Computersystems freigesetzt werden" (Turkle 1988, S. 258).

Gegenüber der Symbolverarbeitung mit ihrer Logik des Funktionierens bzw. Nichtfunktionierens seien dem Neokonnektionismus Konflikt und

Widerspruch (unter seinen Agenten) inhärent. Dabei spielen sog. Zensur-Agenten eine wichtige Rolle. Minsky räumt ihnen zentralen Stellenwert ein, wenn er meint, "daß es keine Intelligenz, künstliche oder andersartige, geben kann ohne Verdrängung" (nach Turkle 1988, S. 259):

"Im Fall der Gesellschaftstheorie ... sind Zensoren immanent. Denn Intelligenz kann ohne Konflikt und Widerspruch nicht entstehen; nur die Gegenwart von Zensoren erlaubt es, daß Intelligenz erwächst. Freuds Beitrag war die Erforschung eines dynamischen Unbewußten. Was unbewußt ist, ist nicht einfach vergessen, alt oder irrelevant für das laufende Funktionieren. Es ist verdrängt. Mächtige Kräfte halten es nieder und das aus gutem Grund. Das trifft für den Computer und die menschliche 'Fleischmaschine' gleichermaßen zu, was verdrängt ist, mußte verdrängt werden - so Minsky" (Turkle 1988, S. 260).

Über die Agententheorie und deren Mechanismus der "kognitiven Verdrängung" konstituierte sich ein "kognitives Unbewußtes" (Minsky): "Das Paradoxe, so Minsky, sei so bedrohlich wie die Urszene" (ebd.).

Das klingt spannend und auch folgerichtig, wenn das Wesen der Verdrängung "nur in der Abweisung und Fernhaltung vom Bewußten besteht" (Freud 1915, S. 108). Noch spannender wird es, wenn wir erfahren, "daß die Verdrängung in der Regel eine *Ersatzbildung* schafft" (S. 114) und "daß die Verdrängung *Symptome* hinterläßt" (ebd.) bzw. daß eine "Wiederkehr des Verdrängten" sich in Symptomen, Ersatzbildungen, Träumen und Fehlleistungen niederschlägt (als die 3. Phase der Verdrängung) (vgl. Laplanche & Pontalis 1986, S. 586).

Demnach schlägt der Neonkonnektionismus eine Maschine vor, die "träumt", "sich verspricht" und eventuell einen hysterischen Anfall produziert - kurz den "neurotischen Computer". Also das können wir beim besten Willen nicht glauben. Vielmehr dient der Mechanismus der "kognitiven Verdrängung" Turkle dazu, uns den Neokonnektionismus als "subversive Theorie" schmackhaft zu machen, indem sie den Begriff der Verdrängung aus seinen (psycho)-dynamischen Zusammenhängen herauslöst und uns die positivistische Version einer "totalen Verdrängung aus gutem Grund" präsentiert. Damit transportiert auch Turkle das von ihr kritisierte Ideal des Funktionierens.

Aber schauen wir uns noch an, welche Auswirkungen das neokonnektionistische Modell auf die "Schnittstelle Mensch-Maschine" (Bammé et al.

1983) hat. Wenn also die Persönlichkeitsstrukturen der Subjekte aus einem maschinisierten und einem genuin menschlichen Anteil bestehen und der Computer den maschinisierten Anteil objektiviert, dann verschwimmt auch hier die Schnittstelle auf Kosten des "irreduziblen Restes" an die Peripherie; allerdings wird dieser uns mit den Etiketten "Subversion", "dezentrales Selbst" und "Fragmentierung" abgesprochen. Könnten wir beim symbolverarbeitenden Modell die Mechanik als "scheinhaft Innerliches" noch der Maschine überlassen (i. S. einer Entmythologisierung eines idealistischen Begriffs geistiger Arbeit) und vermuteten wir, daß sich die Annahme des Unbewußten an sich schon gegen eine Maschinisierung ("Formalisierung und Standardisierung (menschlicher) Austauschverhältnisse und Verkehrsformen" (Baethge & Oberbeck 1986, S. 27)) sperren würde, so belehrt uns der Neokonnektionismus eines besseren. Mit der Konzeption eines "kognitiven Unbewußten" sehen wir uns einer "neuen" Funktionslogik der Maschine ausgesetzt, die unseren psychischen Überlebensstrategien (Verdrängung/Abwehr) ein Maschinenideal aufsteckt (gut verdrängt/schlecht verdrängt).

Wenn wir behaupten, daß über den Mythos des absoluten Funktionierens der Maschine gesellschaftliche Lebensbedingungen erzeugt werden, die primär und sekundär vom Maschinenideal bestimmt sind, dann müssen wir von nun an auch den menschlichen Preis "für die Kulturanforderungen" "Psychopathologie des Alltagslebens" (Freud) am Maschinenideal messen.

Wie schon bei Norman die symbolverarbeitende Maschine, so wird uns auch die neokonnektionistische Maschine als "ideales Objekt" präsentiert. Dies scheint uns ein Charakteristikum der gesamten KI zu sein, das "wesentlich zum Übermenschlichen und Mythischen des Computers" (Löchel & Tietel 1990, S. 101) beiträgt.

Wenn wir uns nun den Anwendungsbereich der (technisch möglichen) KI-Forschung anschauen, den "industriellen Alltag", so wird die ideologische Funktion des "idealen Objektes" - Computer deutlich.

## 5. Der Ernstfall - die industrielle Anwendung der KI

Normans Überlegungen zu den Fehlleistungen beziehen sich unmittelbar auf die Frage, wie in einem Atomkraftwerk "menschliches Versagen" d.h. die Fehlleistung des controllers verhindert werden könne, denn:

"Der Unfall von *Three Mile Island* (d. i. Harrisburg) wurde gerade durch eine (...) Fehlleistung verstärkt. Zwei Ventile im Hilfskühlwassersystem waren wegen Wartungsarbeiten geschlossen worden. Sie wurden nicht wieder geöffnet und die operators benötigen 8 Minuten, um diesen Sachverhalt festzustellen..." (Norman 1980, S. 50).

Die (noch) nicht ganz ersetzbare "unit" Mensch verursacht die Störungen im "unfehlbar" gedachten komplexen System Atomkraftwerk, dem der Mensch nicht mehr folgen kann. Es mutet geradezu rührend an, wenn Norman die Installierung von sog. "forcing functions" (d.s. Kontrollämpchen) vorschlägt, um die Störanfälligkeit der controller-units durch visuelle Reize in Zukunft zu verhindern.

Die "Pathologie" des Systems Atomkraftwerk wird auf den Menschen verschoben; was übrigbleibt, ist dessen "Versagen", dessen Fehlleistung. Wir können uns nur ausmalen, wie nach dem neokonnektionistischen Modell diese "Pathologie" normiert würde. An dem Beispiel kann jedoch deutlich werden, "daß die 'industrielle Pathologie' ... die materialistische Grundlage der 'Psychopathologie des (kapitalistischen) Alltagslebens' (Freud)" (Schneider 1973, S. 222) abgibt.

In diesem Sinn erfüllt das Nachdenken über Fragen wie, ob Computer denken oder ob es eine "subversive KI" gibt, solange ausschließlich ideologische Funktion, als nicht die "industrielle Pathologie", d.h. die Arbeitsbedingungen am Computer, die gesellschaftlichen Bedingungen des Computereinsatzes und dessen Effekte in der Theoriebildung der KI mitgedacht werden.

### Literatur:

- Baethge, M. & Oberbeck, H. (1986). Zukunft der Angestellten. Frankfurt/New York.
- Bammé, A.; Feuerstein, G.; Genth, R.; Holling, E.; Kahle, R. & Kempin, P. (1983). Maschinen-Menschen, Mensch-Maschinen. Grundrisse einer sozialen Beziehung. Reinbek.

- Dörner, D. (1982). Kognitive Prozesse und die Organisation des Handelns. In: Hacker, W.; Volpert, W. & Cranach, N. V. (Hrsg.): Kognitive und motivationale Aspekte der Handlung, S. 26-37. Bern.
- Freud, S. (1901). Zur Psychopathologie des Alltagslebens. Frankfurt/Main.
- Freud, S. (1915). Die Verdrängung. Studienausgabe Band III (1982), S. 103-118. Frankfurt/Main.
- Freud, S. (1915/16). Vorlesungen zur Einführung in die Psychoanalyse. Studienausgabe Band I (1982). Frankfurt/Main.
- Holling, E. & Krempin, P. (1989). Identität, Geist und Maschine. Reinbek.
- Laplanche, J. & Pontalis, J.-B. (1986). Das Vokabular der Psychoanalyse. Frankfurt/Main.
- Löchel, E. & Tietel, E. (1990). Der Computer als evokatorisches Objekt. Psychosozial 13 (3), S. 92-102.
- Neisser, U. (1967). Cognitive Psychology. New York (Deutsch (1974): Kognitive Psychologie). Stuttgart.
- Neumann, O. (1985). Informationsverarbeitung, Künstliche Intelligenz und die Perspektiven der Kognitionspsychologie. In: Neumann, O. (Hrsg.) Perspektiven der Kognitionspsychologie. S. 3-36. Berlin/Heidelberg/New York/Tokio.
- Norman, D. A. (1980). Post-Freudian Slips. Psychology Today (4), S. 42-50.
- Norman, D. A. (1981). Categorization of Action Slips. Psychological Review 88 (1), S. 1-15.
- Schneider, M. (1973). Neurose und Klassenkampf. Reinbek.
- Strube, G. (1990). Neokonnektionismus: Eine neue Basis für die Theorie und Modellierung menschlicher Kognition?. Psychologische Rundschau 41, S. 129-143.
- Turkle, S. (1984). Die Wundermaschine. Der Computer als zweites Ich. Reinbek.
- Turkle, S. (1988). Artificial Intelligence and Psychoanalysis: A new Alliance. Daedalus (1), S. 241-268.
- Volpert, W. (1982). Das Modell der hierarchisch-sequentiellen Handlungsorganisation. In: Hacker, W.; Volpert, W. & Cranach, N. V. (Hrsg.). Kognitive und motivationale Aspekte der Handlung, S. 38-58. Bern.
- Winner, A. (1985). Die "Künstliche-Intelligenz"-Forschung und ihre Auswirkungen auf das Menschenbild in der Wissenschaft. Widerspruch (2), S. 16-28.